CHIRAL GAUGE MODELS ON A LATTICE

Alan Horowitz

Department of Physics, University of Pittsburgh

Pittsburgh, Pennsylvania 15260

email: horowitz@falstaff.phyast.pitt.edu

Abstract: Chiral gauge groups acting on a lattice fermion field are constructed such that all fermion modes (doublers) have the same charge. Details are given for an abelian axial gauge group within a perturbative framework. An action based on this group correctly reproduces the continuum gauge-current anomaly, while preserving global chiral symmetry, locality, rotational symmetry and hermiticity. A Wess-Zumino-like scalar field enters naturally to enforce exact chiral gauge invariance. The degeneracy amongst the doublers can be lifted as in a vector model.

The construction of satisfactory chiral lattice gauge models has long been a thorny problem. Consequently, there is as yet no lattice formulation of electro-weak interactions, even though we've known how to define vector-like models on lattices for a quarter century. It was recognized early on [1] and crystallized into theorems [2] that generic models of free lattice fermions with certain desirable properties, including chiral symmetry, must have equal numbers of right- and left-handed fermions with the same quantum numbers as long as the latter are associated with exactly conserved charges. This is the famous doubling problem. We then expect according to the theorem that left-right asymmetry can be attained if the fermions couple to a non-conserved current. This accords with continuum physics. Indeed in a chiral interaction $j \cdot A$, the divergence of the current j will in general be afflicted by anomalies. On a lattice, this means that the chiral gauge symmetry must be explicitly broken, or equivalently is exact but requires scalar degrees of freedom in the action to cancel the dependence on the gauge parameters. The challenge then is to carefully break the symmetry so that the anomalies come out correctly without sacrificing sacred cows like locality, rotational symmetry and hermiticity. This is what will be done here.

Most approaches to date start with a vector-like model and attempt, with varying degrees of success and clarity, to remove or decouple fermionic degrees of freedom to achieve left-right asymmetry [3-7]. References [8,9] give critical reviews of these and other methods. Perhaps the most promising and scrutinized approach is the Overlap Method [3], inspired by the 5-dimensional domain-wall fermion idea [4]. It is fair to say that this and other methods so far are technically difficult and not easy to analyze. The very recent approach by Lüscher [7] proves the existence of abelian lattice chiral gauge models with exact gauge invariance when gauge anomalies are cancelled. If true this is a significant achievement. Though it is not clear whether this formalism lends itself to practical calculations and how easily it extends to non-abelian models.

The approach given here contrasts with previous ones in that it is based on a chiral gauge group, albeit of an unconventional type, and does not depend on the removal of fermionic degrees of freedom for its chirality. It treats all doublers identically as in a vector model. The doublers thus resemble the families (e, μ, τ) of the standard model. One can then deal with the fermion degeneracy the same way as in a vector model like QCD: a spin-diagonalization to staggered fermions can be made, or Wilson-like terms can be added. It seems possible that the degeneracy can be lifted in such a way that some of the doublers may actually be interpreted as families. Smit constructed models pursuant to this philosophy [10], but there was no handle on gauge invariance.

The formalism will be illustrated on an abelian axial-vector model. The gauge group and the action for the fermions are constructed below as expansions in the coupling constant, though it may be possible to find closed form, non-perturbative expressions. The gauge current anomaly emerges easily as does the gauge-symmetry-restoring Wess-Zumino-

like scalar field familiar from the analysis of anomalous continuum models [11,12]. In previous lattice approaches (excepting the Lüscher formalism) the gauge symmetry is broken more implicitly: the dependence on the scalar field is not easy to extract.

The main motivation of this work is to lay groundwork for the eventual formulation of a lattice gauge model consistent with the phenomenology of electro-weak interactions. For this purpose it may not be a drawback if we could not go beyond a perturbative analysis in the fermion sector. Non-perturbative electro-weak physics seems confined to the Higgs sector. The methodology presented here might lead to insights into how the non-perturbative Higgs sector interacts with the fermions. Also, many people suspect that the chirality of the $SU(2) \times U(1)$ coupling is fundamentally related to the presence of scalar fields, the symmetry breaking, and the massiveness of the vector bosons. Perhaps a lattice approach, such as the one here, can vindicate this suspicion.

To simplify expressions I will use the following notation:

$$K_\mu^{(\pm)}(m,n) \equiv \tfrac{1}{2}\{\bar{\psi}_m \gamma_\mu \psi_n \pm \bar{\psi}_n \gamma_\mu \psi_m\} \tag{1a}$$

$$K_\mu^{5(\pm)}(m,n) \equiv \tfrac{1}{2}\{\bar{\psi}_m \gamma_\mu \gamma_5 \psi_n \pm \bar{\psi}_n \gamma_\mu \gamma_5 \psi_m\}, \tag{1b}$$

where ψ_n is a Dirac fermion at lattice site n. Units are chosen whereby the lattice spacing $a = 1$. For ease of exposition I will take the kinetic part of the action to be that of ordinary naive fermions:

$$S_0 = \sum_{n,\mu} K_\mu^{(-)}(n, n+\mu). \tag{2}$$

As is well known, this action contains 2^d flavors of fermions corresponding to the poles in the propagator at the corners of the Brillouin zone. This degeneracy is a consequence of the invariance of S_0 under "doubling" transformations, $\psi_n \to T\psi_n$, where $T = (-1)^{n_\mu} \gamma_\mu \gamma_5$ or any product of these [13,1]. Including the identity, this is a set of 2^d elements, denoted by $T^{(n)}$. If $\psi_n^{(0)} = u(p)\exp(ipn)$, with $p \approx 0$, satisfies the continuum Dirac equation, then so do the doublers $\psi^{(n)} = T^{(n)}\psi^{(0)}$.

A naive attempt to construct a chiral model, by gauging the axial transformation, $\psi'_n = \exp(ig\gamma_5\theta_n)\psi_n$, is well known to fail. This is most easily seen from doubling transformations of the interaction, $S_1 = ig \sum K_\mu^{5(+)}(m, m+\mu)A_\mu(m)$. For half of the $T^{(n)}$, S_1 changes sign. Thus half the fermionic modes have charge $+g$, while the other half have $-g$. It follows that the model is left-right symmetric. This is not what we want. The problem can be traced to fact that the naive axial transformation does not commute with all the $T^{(n)}$, but only half of them; the rest anticommute.

One of the central points of this letter is that a unitary, axial gauge group can be constructed which commutes with all the $T^{(n)}$ and gives rise to a doubler-symmetric action with the expected continuum behavior. The simplest infinitesimal transformation with the desired property has $\delta\psi_n$ proportional to terms of the form $i\theta\gamma_5\psi_{n+\sigma}$, with $\sigma=(\pm 1, \pm 1, \ldots)$. There are 2 choices for the parameters θ which respect unitarity, namely $\theta_n + \theta_{n+\sigma}$ and

$\theta_{n+\sigma/2}$. Note that in the second choice, θ sits on a site of the dual lattice. I will use this choice because it gives simpler expressions besides being more elegant. It also reveals a kind of duality between vector and axial-vector gauge transformations. (In fact if we omit the γ_5 from the infinitesimal transformation to make it vector-like then half the fermions transforms with positive charge and half with negative, just as in the naive axial transformation of the previous paragraph.) So to first order the axial transformation is

$$\psi'_n = \psi_n + i\bar{g}\gamma_5 \sum_\sigma \theta_{n+\sigma/2}\psi_{n+\sigma} \tag{3}$$

where I have defined $\bar{g} \equiv g/2^d$ for convenience because the sum over σ has 2^d terms. Although this transformation is a matrix in space-time, it commutes with the generators of lattice rotations and translations.

The higher orders in the axial transformation are determined by unitarity and the group composition law. In order to form a group it is necessary that the transformation, starting at $O(g^2)$, involve the gauge field A_μ. (This is reminiscent of, though not directly related to, Fujikawa's derivation of the anomaly [14], wherein the fermionic functional measure becomes dependent on A_μ.) The group property can also be ensured by introducing a group-valued scalar field transforming as $\eta'_n = \eta_n + \theta_n$. However, in the end, the analysis turns out to be simpler using just A_μ, and besides, we want to see how far we can go before we have introduce scalar fields.

Consider 2 successive gauge transformations

$$\psi' = G[\phi, A]\psi$$
$$\psi'' = G[\theta, A']\psi' = G[\theta, A + \Delta\phi]G[\phi, A]\psi.$$

In order to have an abelian group we must have

$$G[\theta, A + \Delta\phi]G[\phi, A] = G[\theta + \phi, A]. \tag{4}$$

This condition, along with unitarity, uniquely determines the $O(g^2)$ term from the $O(g)$ term. Rewriting the transformation, Eq. (3), with this term now included:

$$\psi'_n = \psi_n + i\bar{g}\gamma_5 \sum_\sigma \theta_k \psi_{n+\sigma} \tag{5}$$
$$- \tfrac{1}{2}\bar{g}^2 \sum_{\rho,\sigma} \left\{ \theta_k\theta_l - (\theta_k + \theta_l)[A_{k,l} + \tfrac{1}{2}(\theta_l - \theta_k)] \right\} \psi_{n+\rho+\sigma} + O(g^3)$$

where $k \equiv n + \sigma/2$, $l \equiv n + \sigma + \rho/2$ and $A_{k,l}$ is a sum over As on dual lattice links forming a path between k and l. The path may be taken to minimize the deviation from the straight-line path between k and l. If there is more than one such path, an average is taken. (In Eq. (5) and below, σ and ρ always denote the vectors $(\pm1, \pm1, \ldots)$, while μ

denotes a unit vector.) Unlike the $O(g^2)$ term, the $O(g^3)$ term, which I've calculated, is not uniquely determined by unitarity and Eq. (4). There are 2 free parameters.

To go beyond the perturbative analysis of this paper, it would be necessary to find a unitary, closed form solution of Eq. (4), which reduces to the infinitesimal transformation, Eq. (3). I have not yet made a serious attempt to do this. In such a pursuit, it may or may not be beneficial to recast Eq. (4) as a system of differential equations obtained by taking ϕ infinitesimal and Taylor expanding.

It is easy to see that $\det G = 1$ to all orders in g. This is done using $\det G = \exp(\text{Tr} \ln G)$, expanding $\ln G$ in powers of g, and noting that all odd powers vanish because $\text{Tr} \gamma_5 = 0$ (as does the trace of the space-time matrix at odd powers). This leaves only even powers which being real must also vanish because $\det G$ can only be a unit phase factor, since G is unitary. Since $\det G = 1$, the fermion functional measure is invariant under the chiral transformation, Eq. (5). The anomaly therefore does not appear through the variation of the measure: it appears through the variation of the action.

The local axial gauge group, denoted $\tilde{U}_1^5$, defined by Eqs. (3), (4) and unitarity is not strictly local. At order g^N, G_{mn} has non-vanishing elements out to $|n - m| = 2N$ in 4-dimensions. The fermion action based on $\tilde{U}_1^5$ will thus have long range interactions at higher orders in g. However the $O(g^N)$ term in the action, $\bar{\psi}_m \ldots \psi_n$, with $max|n - m| \approx 2N$, has dimension $4 + N$, and is thus suppressed by a factor of a^N. We therefore expect no violations of locality in processes with energy far below the cutoff, $1/a$.

Now I will construct the action through $O(g^2)$ by applying the gauge principle using $\tilde{U}_1^5$. Starting from S_0 for naive fermions, Eq. (2), the action, invariant to $O(g)$ under the transformation, Eq. (5), is $S_0 + S_1$ with

$$S_1 = -i\bar{g} \sum_{n,\mu,\sigma} K_\mu^{5(+)}(n, n + \mu + \sigma/2) A_\mu(n + \sigma/2) \qquad (6)$$

where the gauge field A_μ sits on links of the dual lattice and has the usual transformation law: $A'_\mu = A_\mu + \Delta_\mu \theta$. In the naive continuum limit, we have the usual axial-vector interaction.

The variation of $S_0 + S_1$ under the transformation can be written after some algebraic reorganization as

$$\delta S_{0+1} = \tfrac{1}{2}\bar{g}^2 \sum \Big\{ K_\mu^{(-)}(n, n + \mu + \rho + \sigma) \big[(\theta_{l+\mu} - \theta_k)^2 - \tfrac{1}{2}(\theta_l - \theta_k)^2 - \tfrac{1}{2}(\theta_{l+\mu} - \theta_{k+\mu})^2$$
$$+ 2\theta_{l+\mu} A_\mu(k) - 2\theta_k A_\mu(l) + (\theta_{k+\mu} + \theta_{l+\mu}) A_{k+\mu, l+\mu} - (\theta_k + \theta_l) A_{k,l} \big] \Big\}$$

where as in Eq. (5), $k \equiv n + \sigma/2$ and $l \equiv n + \sigma + \rho/2$. There is no S_2 depending only on ψ and A_μ which makes δS vanish. The reason is the anomaly. In order to cancel the $(\delta\theta)^2$ terms we need

$$S_2 = -\tfrac{1}{2}\bar{g}^2 \sum_{n,\mu,\rho,\sigma} \big\{ K_\mu^{(-)}(n, n + \mu + \rho + \sigma)[A_{k,l+\mu}^2 - \tfrac{1}{4}A_{k,l}^2 - \tfrac{1}{4}A_{k+\mu,l+\mu}^2] \big\}. \qquad (7)$$

Now with $S = S_{0+1+2}$ we are left with the anomaly in the form

$$\delta S = \bar{g}^2 \sum_{n,\mu,\rho,\sigma} \left\{ K_\mu^{(-)}(n, n+\mu+\rho+\sigma)\left[\theta_k \sum_{P_1} A + \theta_{l+\mu} \sum_{P_2} A\right]\right\} \tag{8}$$

where P_1 (P_2) is an ordered, closed path with vertices k, $l+\mu$, l ($k, k+\mu, l+\mu$).

In order to expose the anomaly in conventional form, we define the effective action, $W[A]$, by $\exp(-W[A]) = \int D\bar{\psi}D\psi \exp(-S)$. It follows that $\delta W[A] = W[A+\Delta\theta] - W[A] = -\ln\langle\exp(-\delta S)\rangle \approx \langle\delta S\rangle$ for small θ, the average being over the fermion fields. In both 2 and 4 dimensions, $\langle\delta S\rangle$ vanishes identically at $O(g^2)$. This agrees with the fact that in 2-dimensions there is no gauge anomaly in the axial-vector model.

To compute $\langle\delta S\rangle$ in 4-dimensions in the continuum limit the following easily derived ingredients are needed:

$$\sum_{P_1} A \approx \sum_{P_2} A \to \tfrac{1}{4}(\rho+\sigma)_\lambda F_{\lambda\mu}$$

$$\langle K_\mu^{(-)}(n, n+\rho+\sigma)\rangle = 2igC\epsilon_{\mu\alpha\beta\nu}(\rho+\sigma)_\alpha F_{\beta\nu} + \dots$$

$$\sum_\sigma \sigma_\mu \sigma_\nu = 16\delta_{\mu\nu}$$

where

$$C = \int_{-\pi}^{\pi} \frac{d^4p}{(2\pi)^4} \sin^2 p_1 \cos^2 p_1 \prod_{j=2}^{4} \cos^4 p_j / (\sum \sin^2 p_\lambda)^2 = \frac{1}{6\pi^2}.$$

Then we find the expected result for 16 flavors,

$$\delta W[A] = -\frac{ig^3}{3\pi^2} \int d^4x\, \theta \epsilon_{\mu\alpha\beta\nu} F_{\mu\alpha} F_{\beta\nu} \tag{9}$$

from which it follows that the current $j_\mu^5 = \partial W/\partial A_\mu$ has the correct anomalous divergence.

Here, j_μ^5 is, of course, the current coupling to the gauge field, and not the current J_μ^5 associated with the usual global chiral symmetry U_A, $[\psi_n' = \exp(i\gamma_5\alpha)\psi_n]$. Without mass or Wilson terms, J_μ^5 remains exactly conserved, keeping the fermions massless. As is well known [1], the conservation of J_μ^5 owes to the cancellation of the U_A anomalies between the doublers. So if we extend this chiral fermion method to the Standard Model, and try to interpret some of the doublers as families, then we could get a different pattern of U_A anomaly cancellation than is currently understood in the continuum model, where the anomalies don't cancel between families. However at this stage this is not an important consideration, since experimental signatures for the cancellation pattern, such as baryon number non-conservation (and B-L conservation) have not been seen, and are not likely to be seen in the near future.

To make this axial-vector model *exactly* gauge-invariant it is not enough to simply add other fermions with appropriate charges to cancel the anomaly. This is apparent from Eq. (8): $\langle \delta S \rangle$ will vanish, but S itself is still gauge-variant. To have exact chiral gauge-invariance, a group-valued scalar field η_n must be introduced transforming as $\eta'_n = \eta_n + \theta_n$. Then inspection of Eq. (8) shows that

$$S = S_{0+1+2} - \bar{g}^2 \sum_{n,\mu,\rho,\sigma} \left\{ K^{(-)}_\mu (n, n+\mu+\rho+\sigma) \left(\eta_k \sum_{P_1} A + \eta_{l+\mu} \sum_{P_2} A \right) \right\} \qquad (10)$$
$$+ \kappa \sum_{n,\mu} \cos[g(\Delta_\mu \eta - A_\mu)] + O(g^3).$$

is gauge-invariant through $O(g^2)$. I added a gauge-invariant kinetic term for η, which is recognized as a fixed-length Higgs term, with η being the phase of the Higgs field. In view of the analysis above, the linear term in η becomes a Wess-Zumino term [15] in the effective action $W[\eta, A_\mu]$.

All mass and degeneracy-lifting terms in this model must be made gauge-invariant with the use of the scalar field η: neither the mass term $\bar{\psi}_n \psi_n$ nor the Wilson term $r\bar{\psi}_n\{\psi_{n+\mu} - 2\psi_n + \psi_{n-\mu}\}$ is gauge-invariant by itself: gauge-invariance requires Yukawa interactions as in the Standard Model. It is straightforward, using the chiral gauge transformation, Eq. (5), to construct these interactions. For example, the gauge-invariant mass term to $O(g^2)$ is

$$\bar{\psi}_n \psi_n - 2i\bar{g} \sum_\sigma \bar{\psi}_n \gamma_5 \psi_{n+\sigma} \eta_{n+\sigma/2} - 2\bar{g}^2 \sum_{\sigma,\rho} \bar{\psi}_n \psi_{n+\sigma+\rho} \eta_{n+\sigma/2} \eta_{n+\sigma+\rho/2} \qquad (11)$$

which reduces in the naive continuum limit to the expansion of $\bar{\psi}_x \exp(-2ig\gamma_5\eta_x)\psi_x$ to $O(g^2)$. It is interesting to note that this lattice mass term is not invariant under the U_A global chiral symmetry, while its naive continuum limit is (provided $\eta \to \eta + \theta$ under U_A).

Because the action is doubler-symmetric, we can reduce the number of fermion modes by spin-diagonalizing in the usual way [16]: a unitary transformation, $\psi_n \to \gamma_1^{n_1} \ldots \gamma_4^{n_4} \psi_n$, breaks the action into 4 identical, independent pieces. Tossing 3 pieces away we have an action describing 4 Dirac fermions (reduced from 16), with one fermionic degree of freedom per lattice site (reduced from 4). Then terms can be added to lift the mass degeneracy of these fermions. Interestingly, The interaction S_1, linear in A_μ, in terms of these staggered fermions is identical to the linear term in Smit's staggered fermion approach [10] except for the coordinates of the gauge fields. Higher order terms are however not similar: Smit's action has no gauge symmetry.

How should we use the action, Eq. (10) ? Anomalous models, such as in Eq. (10) are believed to be nonrenormalizable [see, e.g. ref. 12]. To obtain a renormalizable continuum model, the gauge anomaly should be cancelled by other fermions, as for example in a model with 9 fermion fields, 8 with axial charge, $q = 1$, and one with $q = -2$. To extract physics from such a model, the parameters must be near a critical point or surface. Ignoring

the fermions for the moment, the Higgs model with just η and A_μ has a Higgs phase for $\kappa > \kappa_c$ with $M_A > 0$ and a symmetric phase for $\kappa < \kappa_c$ with $M_A = 0$. With an eye towards electro-weak physics, let's say we want to obtain a continuum limit in the Higgs phase with a 3 parameter model, κ, g and y, the coefficient of the mass-Yukawa term in Eq. (11). At a point in the Higgs phase near the critical surface, for small g and y the truncation of the action at some low order in g should be valid. To calculate some observable of ψ or A_μ one could either perturbatively integrate out the fermions in the path integral and numerically integrate (simulate) the resulting expression in A_μ and η, or include the fermions in the simulation. One would have to check that answers were not sensitive to the order of truncation of the action.

As we get closer to the critical surface, the fermion mass m_f, in units of inverse lattice spacing, should approach zero (otherwise there would be no low mass fermion in physical units). To get $m_f \to 0$ we can either tune y to zero as $\kappa \to \kappa_c$ or keep y fixed and rely on the higher orders in Eq. (11) to reduce m_f. At larger values of y we will need more terms in Eq. (11). Similar reasoning might apply to an attempt to remove doublers using a Wilson-Yukawa term: the fluctuations in η may conspire to send the effective amplitude of such a term to zero, giving the doublers a mass much less than the cut-off. This is what happens in the Smit-Swift model [17,9]. Unlike in the Smit-Swift model the chirality of the model here does not depend on the Wilson-Yukawa term. If need be, the number of doublers can be reduced by using staggered fermions or other methods.

The extension of the above methodology to a V-A abelian model is straightforward. We want a transformation, which in the naive continuum limit reduces to $\psi'(x) = \exp[ig\theta_x \frac{1}{2}(1-\gamma_5)]\psi(x)$, and which transforms all doublers the same. This is accomplished by the following infinitesimal transformation,

$$\psi'_n = (1 + \frac{i}{2}g\theta_n)\psi_n - \frac{i}{4}\bar{g}\gamma_5 \sum_\rho (\theta_n + \theta_{n+\rho})\psi_{n+\rho}.$$

The higher orders are constructed using the group property, Eq. (4), and unitarity as was done above for the axial transformation. The extension to non-abelian chiral groups is slightly more subtle, and will be presented in another publication.

[1] L.H. Karsten and J. Smit, Nucl. Phys. B183, 103 (1981).
[2] H.B. Nielsen and M. Ninomiya, Nucl. Phys. B185, 20 (1981); Nucl. Phys. B193, 173 (1981).
[3] R. Narayanan and H. Neuberger, Nucl. Phys. B443, 305 (1995).
[4] D.B. Kaplan, Phys. Lett. B288, 342 (1992).
[5] A. Borelli, L. Maiani, G.-C. Rossi, R. Sisto and M. Testa, Nucl. Phys. B333, 335 (1990).

[6] J.L. Alonso, Phys. Rev. D44, 3258 (1991).

[7] M. Lüscher, hep-lat/9811032.

[8] Y. Shamir, Nucl. Phys. (Proc. Suppl.) B47, 212 (1996).

[9] D.N. Petcher, Nucl. Phys. (Proc. Suppl.) B30, 50 (1993).

[10] J. Smit, Nucl. Phys. (Proc. Suppl.) B4, 451 (1988); 29B,C, 83 (1992).

[11] E. D'Hoker and E. Fahri, Nucl. Phys. B248, 59 (1984); L.D. Fadeev and S.L. Shatashvili, Phys. Lett. B167, 225 (1986).

[12] R.D. Ball, Phys. Rep. 182, 1 (1989).

[13] A. Chodos and J.B. Healy, Nucl. Phys. B127, 426 (1977).

[14] K. Fujikawa, Phys. Rev. Lett. 42, 1195 (1979).

[15] J. Wess and B. Zumino, Phys. Lett. B37, 95 (1971).

[16] N. Kawamoto and J. Smit, Nucl. Phys. B192, 100 (1981).

[17] P. Swift, Phys. Lett. B145, 256 (1984); J. Smit, Acta. Phys. Pol. B17, 531 (1986).